

ESO



PATENT ABSTRACTS OF JAPAN

(11) Publication number: 11328218 A

(43) Date of publication of application: 30.11.99

(51) Int. Cl. G06F 17/30
G06F 12/00

(21) Application number: 10146539

(22) Date of filing: 12.05.98

(71) Applicant: NIPPON TELEGR & TELEPH
CORP <NTT>

(72) Inventor: HIKITA TOMOHARU
MATSUMOTO MASAOKI
FUJISHIRO NORIKO

(54) CONTENTS ATTRIBUTE INFORMATION
NORMALIZATION METHOD, INFORMATION
COLLECTING/SERVICE PROVIDING SYSTEM,
ATTRIBUTE INFORMATION SETTING DEVICE
AND PROGRAM STORAGE RECORDING MEDIUM

form suited to a retrieval service or the like. Further,
as of the contents attribute information whose the
structure is normalized, a character expression form is
normalized and numerical expression is normalized.

COPYRIGHT: (C)1999,JPO

(57) Abstract:

PROBLEM TO BE SOLVED: To construct a data base for
providing a service without being restricted by the
structure/form of a document for the perusal of an
information provider by performing normalization
processing of an attribute structure as of contents
attribute information and performing normalization of a
character expression form and normalization processing
of numerical expression.

SOLUTION: An automatic information collecting part 101
of an automatic information collecting and classifying
device 100 patrols a Web site 120 of an information
provider on a network 110, collects document files and
extracts information. An attribute extraction part 102
normalizes a character code and then extracts only the
contents attribute information as of the collected and
extracted documents. An attribute normalization part 103
refers to a normalization rule 106 and normalizes the
contents attribute information provided with the
structure/form coincidently with a perusal document
extracted by the attribute extraction part 102 to the

